Multicore Computing and the Future of Seismic Imaging

Michael P. Perrone *(IBM Research)*

With some of the largest supercomputing clusters on the planet, the seismic imaging industry has a tremendous appetite for computing resources. Increasing energy demand and societal pressures for greener energy will only accelerate this growing need. It is therefore clear that seismic imaging companies will be early adopters for the newest generation of supercomputers that enable petascale computing – machines that are perhaps only a few months away. However, as this transition to petascale occurs, the industry will also be one of the first groups to come face-to-face with the significant changes that will be required in order to go beyond petascale. The industry must actively develop plans for how their business models and standard operations will change in order for practitioners to achieve the performance they require for their ever-growing technical challenges.

For decades, we have relied mainly on the increasing clock speeds of “traditional” microprocessor architectures for computational performance gains. However, in recent years this approach has been hindered by the physical limitations of semiconductors and by traditional processor architecture implementations. Issues with power consumption, heat dissipation and memory latencies have led to diminishing returns on performance. High performance computing applications such as digital content creation, electronic design automation, image and signal processing, financial algorithms, scientific research, and seismic processing, may need a fundamentally new technology and approach to the system-level architecture to achieve the desired level of performance.

The next few years will present very important challenges and opportunities for high performance computing. Petascale computing and, eventually, “exascale” computing will bring the promise of the capability to deliver more complete solutions to some of the most challenging and complex issues facing the industry. However for well documented technology reasons, these new computing systems architectures will be radically different in design from traditional high performance computing platforms.

For example, in response to growing technological obstacles, the processor industry is moving down the multicore path. This development is driving a sea change in the computer industry for which a new "Moore's Law" may be arising - dictating a doubling of the number of cores per unit time. As more cores are available per chip, the old programming approaches will not be adequate to achieve the performance required by this industry; and naive assumptions of linear scaling of performance with the number of cores will be very wrong.

Recent experience with multicore has identified key challenges which will have to be overcome in order to realize the potential of the next generation of supercomputing. These challenges include fundamental algorithm design; integration of novel architectures with more traditional computational systems; management of the unprecedented amounts of data which are now a key component in all high performance computing activities; and the development, improvement and validation of new applications solutions which address the full complexity of the problems which these novel architectures will make tractable.

To be truly successful, the industry will need to adapt and change its current computational practices to prepare for the impact of these looming changes. The key

will be a laser-like focus on what practitioners can do to achieve the full potential of multicore-based petascale supercomputers and eventually exascale computing.

High-Performance Multicore for Seismic Imaging

One recent addition to the race for sustained petaflop performance is the IBM BladeCenter QS22 blade server. Each QS22 features two multicore 3.2GHz IBM PowerXCell 8i processors with up to 32 GB of memory, 51.2GB/s of main memory bandwidth, dual Gigabit Ethernet and an optional dual-port 4x InfiniBand HCA. With over 400 GFLOPS of single precision peak performance per blade, the QS22 provides extraordinary computing density—over 22.4 TFLOPS peak performance for a rack of blades. Additionally, the QS22 has over 200 GFLOPS of double precision peak performance. These features make Cell ideally suited for seismic image processing.

The PowerXCell 8i processor on the QS22 has 9 heterogeneous cores connected by a 288GB/s on-chip bus and has been designed to be very power efficient. The PowerXCell 8i processor is an asymmetric multi-core processor that is optimized for parallel processing and streaming applications. Unlike symmetric multi-core, cache-based architectures which may not be able to efficiently handle streaming applications, the PowerXCell 8i processor is designed to offer very high performance and fast response. The PowerXCell 8i processor includes a Power Processor Element (PPE) and eight highly optimized enhanced double precision (eDP) SIMD engines called Synergistic Processor Elements (SPE).

PowerXCell 8i processor performance is about an order of magnitude better than traditional processors for media and other applications that can take advantage of its SIMD capability. The PPE is intended to run the operating system and coordinate computation. Each SPE is able to perform mostly the same as, or better than, a General Purpose Processor (GPP) with SIMD running at the same frequency. A key performance advantage comes from its eight de-coupled eDP SPE SIMD engines with dedicated resources including large register files and DMA channels.

With a peak of over 2 GFLOP/Watt, the PowerXCell 8i is an excellent solution for energy starved data centers. This means that in today's energy constrained environment, the QS22 may allow significant datacenter energy savings and thereby reduce the total cost of ownership. And because the overall BladeCenter infrastructure uses super energy-efficient components and shared infrastructure architecture, you can realize even lower power consumption when compared to many alternative designs.

An additional benefit of QS22 blades is that they can be intermixed in the BladeCenter H chassis. Intermixing allows users to design clusters with the optimal balance of computational resources for their workloads.

Cell-Based Supercomputers

Several groups in the high-performance computing community have already seen the potential of Cell for supercomputing and have begun Cell projects. The most well-know is possibly the **Roadrunner Project** at Los Alamos National Labs (http://www.lanl.gov/roadrunner) which is targeted to produce a machine with a peak performance of 1.6 petaflops and a sustained performance 1.0 petaflops. It will be a hybrid system of 8000 Cell blades and 4000 Opteron blades targeted to be completed this year.

Another Cell-based project which is much more relevant to the seismic community is the Repsol **Kaleidoscope Project** in collaboration with the Barcelona Supercomputing Center, 3DGeo and CSIC (http://www.bsc.es/projects/kaleidoskope_tmp). This system is envisioned as a pure QS22 cluster in the near-term and evolving to the next generation of Cell processors in the 2010 time frame. The goal of this project is nothing less than to revolutionize the way the industry thinks about seismic imaging. The project is based on the premise that the superior performance of the Cell processor will enable a new generation of algorithms that will enable faster and more accurate seismic imaging which will, in turn, allow practitioners to focus their expertise on finding oil and will enhance their value to the industry.

Why Cell for Seismic?

On a chip-to-chip basis, the QS22 can be an order of magnitude or more faster than existing solutions for some of the critical workloads in the industry, such as Fast Fourier Transforms (FFT), Wave Equation Migration (WEM), Reverse Time Migration (RTM) and the Inverse Multiple Attenuation (IMA) algorithm for the Mission-Oriented Seismic Research Project at the University of Houston. It is this kind of game-changing performance increase that practitioners need to take advantage of in order to improve imaging quality and speed.

The reason the PowerXCell 8i processor can deliver superior performance is a result of the additional programmer control that it provides over the flow of data. Historically, computer science algorithmic improvements have been driven by a desire to reduce the number of operation required for a given calculation because it was the computations that typically took the majority of the time. However due to a variety of technological constraints, the balance between computation and memory access is shifting – particularly for multicore processors. As processors achieve increasingly high GFLOP rates, the number of operations per processor clock cycle has increased beyond the FLOPs per byte ratio of many important workloads. This can be see most notable in Reverse Time Migration where in many cases the algorithm is bandwidth-bound. And this trend will continue as processors include more an more cores. Grow in computation per processor will continue to outstrip growth in bandwidth – making the problem even worse.

This trend has led many in the field to realize that we are headed towards a major shift in programming style in the multicore future. Mike Acton of Insomniac Games, a video game manufacturer that develops high-performance code for multicore platforms, likes to say, “It’s all about the data!” By this he means that high performance will only be achieved if practitioners are willing to modify their algorithms to put data-flow management front and center. algorithms that do not specifically adapt to make data handling their primary focus

One of the wonderful things about the Cell processor is that it designed specifically to allow the programmer to have complete control of the data flow on the processor. Cell programmers can make certain that the data is where they want it when they want it. It is this detailed level of control that allows the Cell to efficiently use its bandwidth and thereby to achieve significant percentages of peak performance – typically much higher than other processors.

In the same breath, one must state that this control is also one of the most highly criticized aspects of Cell. This high level of data flow control comes at the cost of more detailed – and many claim more difficult – programming. There is no doubt that today's Cell tool chain puts more burden on programmers; but it is just this additional burden that allows the Cell to perform so well.

It is well known that many high-performance workloads do not scale linearly with additional cores on multicore processors. For example, you might expect to get a 2x performance improvement when going from one core to two cores on a chip, when in practice you might only get 1.7x. Similarly, going from two cores to four generally won't even give 1.7x. As this trend continues to 32 and 64 cores and beyond, the problem just gets worse. In the end, you will have many cores most of which are not operating anywhere near peak performance. And the end result? A very hot data center with modest increases in performance. This is not a sustainable path for the industry.

The reason for this poor scaling is that other multicore processors do not allow the same level of data control that one has on the Cell processor. One might be able to modify algorithms for other multicore platforms; but increasingly, people are beginning to realize that the amount of effort required to tightly-tune other multicore processors is actually more than the effort required by Cell. Recent work by Sam Williams at the University of California, Berkeley (http://www.cs.berkeley.edu/~samw/) highlights this issue by the comparison of the optimization of several high performance computing workloads on a variety of multicore processors.

The bottom line of these results is the growing realization that in order to achieve the full benefits of multicore architectures, practitioners will have to seriously re-think the implementation of their algorithms in order to guarantee that sufficient attention is given to the tight management of data flow at the processor level. The Cell processor is the first multicore processors to make detailed data flow explicit, compulsory and comparatively easy.

In essence, the Cell processor architecture forces the programmer to face the multicore challenge head-on with the tools and the flexibility to do it properly and efficiently. This is the true advantage of Cell.

How Does Cell Help?

Let's look at some specific, common, seismic workloads to see examples of how the detailed data flow control available on the Cell processor facilitates higher performance.

Traditional processors rely on caching to enhance data access from main memory; however if the data set does not fit in cache, these processors suffer very large latencies to access the data that doesn't fit. Furthermore, if the data is accessed in unpredictable patterns, the performance can suffer dramatically due to frequent cache misses. This problem exists in many seismic workloads.

In the case of 2-dimensional Fast Fourier Transforms, a common data size of about 1000x1500 complex, single-precision elements requires 12MB of memory. This size

is bigger than most of today's caches and is continuing to increase as seismic surveys grow in size. So cache misses are clearly a problem and the migration to multicore only adds more stress to an already over-burdened cache. Compounding that problem is the fact that FFT data access patterns are very irregular which leads to frequent requests for data that is not in cache, even when the data set is only slightly larger than the cache.

In the case of Wave Equation Migration, the velocity model can be 1GB or more and will clearly not fit in any processor's cache. However, the Green's function that is used to propagate the wave depends on the velocity field and must be read for each point in the velocity model. This can lead to very irregular data access of main memory with a corresponding drop in performance.

In the case of Reverse Time Migration multiple gigabyte data sets must be repeatedly brought in and out of the processor in an operation that requires very few FLOPs per byte due to typically very sparse stencils. So a straightforward port is very wasteful of bandwidth. It is critical that the data flows in RTM be orchestrated very carefully to get efficient usage of the processor.

In each of the cases described above, the Cell processor provides the control and flexibility to guarantee that the data is where you want it when you want it. This level of control allows the user to hide the latency of memory transactions behind the computations as well as to make sure that the bandwidth is efficiently used without repeated accesses of the same data unnecessarily. This level of control enables the efficient and atomic movement of data that leads to high performance.

It's All About the Data

We are at the cusp of a potentially radical change in the seismic imaging industry: Computational power is now available that could reduce imaging times from weeks and months to an hour or less. How would business processes change if imaging could be multiple times a day? How much more productive would geophysicist be? How would their workflow change? How much more oil would be found?

In order to get to achieve this game-changing, the industry will have to rethink its algorithms and how they are implemented. The precise and efficient flow of data will take center stage and the algorithms and hardware that enable that level of control will become critical.